Peer Review Report for

"Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data"

Reviewed by
Dr. Alicia Carriquiry, Iowa State University
Dr. Jean D. Opsomer, Colorado State University
Dr. Stephanie Eckman, Institute for Employment Research, Nuremberg, Germany

Introduction

This document combines the comments provided by three different peer reviewers of the MRIP Project Report entitled "Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data." The document provides verbatim reviewer comments without identifying the source of each comment.

Reviewer 1

In this report, an approach for estimating charter-boat catch and related characteristics based on logbook data is proposed. The approach uses Bayesian modeling, making it convenient to specify separate models for number of trips, average effort per trip and average catch per unit effort, and combining their estimates and associated measures of precision into a single estimation procedure. The approach is demonstrated on two species, red snapper and vermillion snapper. Finally, the report gives a set of recommendations and conclusions. The overall estimation approach is carefully described and statistically valid, and there is a large literature on implementing Bayesian models such as those proposed here.

While the approach is statistically valid and appropriate, I have three somewhat interrelated concerns regarding its implementation in the context of NOAA's overall goal of producing official estimates of catch and related characteristics from these data. I will describe those first, followed by further discussion, and I will number my comments for ease of reference.

1. The first and most important concern has to do with the non-random selection of the logbook data. This was recognized early on, when it was clear that the logbook records could not be treated as a census of trips. However, they are treated here as a random sample of trips, which is better than a census but which still might not be appropriate, because of non-random selection of reported trips. Looking specifically at the three identified components of catch, I see the estimation of the first one (number of trips) as being fine, but I am not sure about the remaining two, because they assume that the logbook trips are a representative sample of all trips in the estimation procedure. I realize the estimates are being compared with those obtained from dockside sampling in this report. But the proposed method does not actually provide any way to incorporate the dockside sampling estimates, and they are only used as an external check to detect potential large discrepancies. I will return to this issue in comments (4) and (5) below.

2. The second concern has to do with the fact that, being model-based, appropriate models need to be picked for each of the components of the estimation procedure. This needs to be potentially repeated for every variable (catch vs. released), species, each time period, etc. This was illustrated in this report with the change in the models between red and vermillion snapper, for instance. Doing this properly for the range of estimates that are to be produced is a very significant amount of work, which will need to be done and documented to justify the estimates. In general, such modeling effort is usually only undertaken if design- based methods cannot be used, for instance because the sample sizes are too small or because issues such as informativeness, nonresponse or measurement errors are present to such a degree that they need to be adjusted explicitly via modeling. It is not clear to me that the logbook data fall in either of those categories. Let me reiterate that I do not think the proposed estimation method is not statistically valid or otherwise incorrect. The issue is that implementing this for the official for-hire catch estimation will entail a very substantial amount of work.

3. The third concern relates to the measures of uncertainty that are produced as part of the Bayesian estimation. As explained on p.30 of the report, the Bayesian estimation approach provides these measures of uncertainty as part of the estimation procedure, which is a indeed major strength of the approach. However, these measures of uncertainty are valid under the critical assumptions that (1) the parametric models for the various components of the final estimator are correct, and that (2) the observed and unobserved parts of the target population follow the same distribution (i.e. that the nonrandom selection of the observed units did in fact produce a representative sample). Whether these assumptions are met or not has a major effect on the precision and reliability of the estimates, so the concern is therefore that the precision of the estimates as produced by the posterior distributions could be overstated, potentially severely so.

4. Returning to the first concern, I want to discuss the use of the dockside sampling data in validation and possibly estimation of the for-hire catch. As done in this report, these data can be used to compare with the estimates coming from the logbook data to assess informativeness of the latter, as was done in this report. However, what I would have preferred is not to compare the model fits using both data, but to compare the model fits using the logbook data to \un-modeled" (design-based) estimates from the dockside data. While having both model fits agree is certainly a good sign, there is the potential for both being wrong in the same direction if the assumed model is not appropriate. Note that I think this is unlikely for these species, which were carefully modeled, but as a recommendation for future comparisons, I think that comparing a model prediction with a model-free estimate is a better approach.

5. But in addition to using the dockside data for comparison, I am wondering whether a procedure that incorporates them in estimation might not be preferable. Ideally, I would like to be able use the logbook data to estimate the average effort per trip and average catch per unit effort for the logbook trips, and the dockside data to estimate the same quantities for the

non-logbook trips. This would correspond to splitting the population in two components, with a census of one (logbook trips) and a random sample of the latter (non-logbook trips). Another possibility is to treat this as a dual-frame problem, with the dockside sampling reaching both the logbook and the non-logbook trips. This would completely remove the selection bias concerns, but I realize both of these approaches might be very difficult to do in practice, because they require a determination of the \logbook status" of individual dockside observations. If that is not feasible, I still think it would be worthwhile to investigate ways to incorporate the dockside data directly into the overall estimation procedure, possibly by calibrating relevant quantities in the logbook estimation procedure to the dockside results. The goal here is to remove potential sources of informativeness to the extent possible. As a simple example of this, one could compute the estimated average number of anglers and/or the average trip length for the dockside data, and use those to calibrate the corresponding quantities for the logbook data. I am not sure whether this can be incorporated into a Bayesian estimation framework, but conceptually, this would remove potential bias in the logbook data that would be caused by the logbook trips being either of different size or of different length from the non-logbook data. Even if the dockside data are not used in estimation, a set of relevant measures of comparison between both the dockside and the logbook data should be developed and tracked over time, to ensure (and confirm to the data users community) that the logbook data are representative of the overall for-hire catch.

6. An important issue identified in several places in the report is that of data review and editing. The presence of outliers, such as those discussed in the report, clearly has a large effect on the resulting estimates of catch. This will need to be addressed regardless of the estimation method eventually adopted.

7. Finally, I found a number of minor typos that should be fixed: “allude" instead of “elude" (p.19), “available" not “avaliable" (p.21), “amount" not “amont" (p.28).

Reviewer 2

The author has written a report proposing a new technique to estimate catch, release and mortality of fish species in the Gulf coast of Florida. The report is very well and clearly written. Bayesian estimation techniques are not easy to describe in non-technical language, and the author has done a very good job with the task. I am also convinced, based on the material available in this report, that the estimation technique is a sound one, though the report would be more persuasive if it also included a comparison to more standard estimation methods. I have a few comments on how the text could be further improved.

General Comments:

I strongly recommend that the report include a paragraph or two in the introduction on the use and importance of estimates on catches and releases from for-hire boats. While it is obvious to the author why this data is needed, it is not necessarily obvious to readers and the report will be stronger if it (briefly) makes a case that this data is crucial to government and industry.

The report repeatedly gives 95% credible intervals but does not explain how such intervals are to be interpreted. One or two examples in the report of how to interpret such intervals should be added.

The report should include an outline to let the reader know how it is structured.

Please use commas in numbers with more than 3 digits before the decimal point – this will make the text and tables easier to read.

I was expecting to read in the conclusion some recommendation about whether collecting dockside data was worth the cost. Does the author feel that the proposed technique using only the logbook data is sufficiently robust that the dockside data is not needed? Would the logbook data be as good as it is if the dockside data were not collected alongside? (That is, might boat operators give lower quality data if they know that no dockside validation will be done?)

Specific Comments:

Page 3 – I would like to see a bit more discussion of the results of the MRAG Jan 2006 report that found the estimation approach in this report performs better than a survey based estimator.

Page 11 – spell out in words what the values in the w and eta vectors are. I believe w_j is the number of trips where the cpue fell in the the jth bin and eta_j is the proportion of trips falling into this bin.

Page 13 – similarly, more words are needed here about what the values of y_i and w_j are and what n and k are.

Page 14 – please don't use m here to index the number of simulations, as m is used on page 12 and elsewhere as the number of verified trips.

Page 15 – here, or perhaps above, a discussion is needed about why we want M_1 different values of the total catch. The reader may be worried that the algorithm produces not a single number but a large vector of numbers. Explain that one can use this vector of values to get a sense of the distribution and make statements about the estimate's precision.

Page 16 – first full paragraph, last sentence. The use of "information on catch" and "catch information" confused me. Please use the same phrase in both parts of the sentence to emphasize that the important difference is between logbook and dockside sources. I also suggest italicizing logbook and dockside to draw the reader's attention to the fact that this is the important distinction you're making.

Page 17 – point b. Above you recommended using a diffuse distribution here as prior. You might make that point again here.

Page 18 – point 2. When you give the value of 50,000 please indicate that this is M_1 in the algorithm.

Page 19 – point a. M_1 is also 50,000 here, but it needn't be the same as above, right?

point b. again, include references to the notation used in other parts of the text. "For each pair of values (\alpha^*_m, \beta^*_m) simulated from the posterior, M_2 = 10,000 y^*_{m,q} values were then simulated from the data model (7) using those values as parameters."

As above, though, avoid the use of m as an index.

point c. here you do a nice job referring back to the earlier notation.

Page 20 – please add a point 5 which states that you multiplied the three estimates together to get tau.

Page 25 – first full paragraph. This paragraph needs a topic sentence to make it clear to the reader what the point is. The last sentence starts with "the point is" but by then the reader is far too lost. The first sentence of the paragraph is about red snapper, but the rest of the paragraph is not. Please rewrite this paragraph.

Also, the text talks about % of trips with any vermillion snapper caught, but the table presents only numbers. Percents seem much more relevant to the discussion, so redo the table in terms of percents.

Conclude this paragraph with a statement that the next section will address how the high proportion of trips without vermillion snapper catches necessitates a change in the estimator.

Page 26 – first line. Refer the reader back not only to section 3 but also to a specific step in the algorithm.

Page 27 – why do you give such rough numbers here? You have better numbers in the table and numbers in the text that match the tables would help the reader see where you're pulling the numbers from. Also, 250,000/12,000 is more than twice as many; it's 20 times as many. Or does the "about twice" in the text refer to a different ratio?

Page 28 – missing ")" near the end of the page. Amount misspelled as amont. "there appear to be groups" is very vague. Trends would be a better word here.

Page 30 – first paragraph. This paragraph is quite defensive. The author seems to be arguing against critics that are not known to the reader. Either cite some articles or correspondence from these critics or tone down the defensiveness in this paragraph.

point 1 – this point also seems to be arguing against a suggestion that the reader has no knowledge of. What kind of size eligibility criteria are currently used for the logbook study? What does the author propose for those criteria? In contrast to the rest of the report, this section is not clear.

Page 31 – point 5. The reference here should be to Figure 12, not Figure 9.

Page 33 – "can do" – something is wrong in this sentence.

Page 34 – top of page. Again the author is responding to criticisms that the reader is not aware of. First tell us what these criticisms are, with references or quotes, and then rebut them.

Towards the end of this page there are a few typos.

Tables and Figures

Table 1: the text mentions that the first three columns are m, X and M, but these labels should be included in the table.

Table 6 (and all others): include commas in numbers.

Table 11: make "With Harvest" columns into %s, not numbers

Figure 2 (and others): include kappa in the horizontal axis label to make it clear to your reader what is represented.

Figure 4: again, label top horizontal axis as theta hat posterior and bottom horizontal axis as Nhat

Figure 6. The horizontal axis here is kappa*

Reviewer 3

I enjoyed reading Mark Kaiser's report on estimation using logbook data. Kaiser does a very good job of motivating the methods and explaining assumptions and limitations of the estimators he proposes for the various quantities of interest.

At the outset, I will say that I fully agree with a conclusion in this report. Using logbooks, at least according to the information obtained from the pilot study, appears to be a reasonable sampling strategy, which, in conjunction with the appropriate statistical methodology will almost certainly result in more accurate estimates of catch, release and other totals than a poorly conducted census.

There is little discussion in the report about whether everyone is expected to fill the logbook in every trip or whether at some point a more formal sampling approach might be conducted to select a subset of the trips for logbook completion. This point, which can be discussed at a later time, in no way would change the methods or the conclusions that are drawn in this report.

If there is a limitation to this methodology, it is that it requires some expertise by the person in charge of producing the estimates. This type of analysis is difficult (but not impossible) to carry out using "canned software", so at least initially, MRAG might need to depend on a consultant to carry out analyses. One other limitation that received little attention in Kaiser's report is that for some of the parameters (e.g., cpue), it might be necessary to develop a different model for each different species. As a consequence, it is probably not possible to write a "black box" type of program that MRAG can use on all types of fish. I revisit this issue briefly in my specific comments.

Overall, I find that the methods proposed in the Kaiser report are correct, appropriate and well justified and that Kaiser's approach offers an excellent alternative to the methods that appear to be in use today. My specific comments are mostly picky and offered as suggestions for the future version of the methodology

Specific comments

- On page 5, something might be said about other choices for the prior parameters α,β. How sensitive are the results (given the data) to different choices of these two parameters? Probably not terribly sensitive because of the large sample sizes, but even a small change in the posterior mean of θ could have a noticeable effect on τ because of the multiplicative structure of (1).

- Page 7, equation (7): why not estimate δ together with the other parameters in the model? Perhaps I am missing something here, but it does not seem like including δ in the MCMC process would be too difficult. If nothing else, at least it would be good to understand why the choice of $\delta = 0.5$ and also get a sense of the sensitivity of the estimates of this (different) set of parameters denoted α,β to different values of δ.

- Page 10, toward the top: I very much appreciate the difficulties with modeling a variable such as cpue. One alternative that is not mentioned is to use a semi-parametric mixture approach. The binning approach proposed by Kaiser is fine, but the bins are arbitrary and would need to be revisited every time a new dataset is analyzed. Mixtures, in contrast, can be formulated so that the number of components in the mixture is estimated from the data. Without going to such a "high tech" extreme, a simpler model in which the number of components is limited to, e.g., three, ought to be flexible enough to account for almost any shape in the distribution of cpue. One advantage of using a mixture model is that developing "black box" software is easier.

- I missed mention of δ in the applications to red and vermilion snapper data. Was it fixed at 0.5 always?

- Page 30, point 1: I do not know what proportion of fishing is done in small recreational boats versus large vessels. Kaiser is correct in that mixing apples and oranges is typically a bad statistical idea. However, rather than eliminating the large fishing vessels from the dataset, I would probably carry out the analysis for the two types of boats separately and then combine the results using some kind of weighting that depends on the relative contribution to catch by the different types of boats.

- Page 31, point 9: The scatter plots clearly show a non-linear association between effort and cpue. What if cpue were transformed (using, e.g., a square root transformation). Would the association become more linear?

- Page 32, first full paragraph: I agree with the points raised by Kaiser, but a question is why not simulate the two variables using a bivariate proposal instead of assuming independence. If a transformation could be found to linearize the association between the two variables, carrying out a bivariate simulation would not be difficult.